

### **REMARKS**

Claims 1-19 are pending in this application, of which claims 1, 17, and 18 are independent in form. Reconsideration of the non-final Office Action dated April 13, 2009 is respectfully requested in view of the foregoing amendments and the following remarks.

#### **35 U.S.C. § 112 Rejections**

Claims 1, 2, 4, 8, 10, 11, 17, 18, and 19 were rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the written description requirement. In particular, the Examiner stated (April 13, 2009, office action, page 4):

The claim(s) contains subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention. The term "specification" is not defined in the specification as for an individual to understand the metes and bounds of the claim language.

The applicant respectfully disagrees. However, for the purpose of advancing prosecution, claims 1, 2, 4, 8, 10, 11, 17, and 18 have been amended to remove recitation of a "specification." Withdrawal of the 35 U.S.C. § 112, first paragraph rejection of the claims is requested.

#### **35 U.S.C. § 101 Rejections**

Claims 1-3, 5-16, and 18-19 were rejected under 35 U.S.C. 101 as being directed to non-statutory subject matter. The Applicant has made certain amendments to independent claims 1 and 18 to advance prosecution. Withdrawal of the 35 U.S.C. § 101 rejection of the claims is requested.

#### **35 U.S.C. § 103 Rejections**

Claims 1-19 were rejected under 35 U.S.C. 103(a) as being unpatentable over US5,797,123 (Chou) in view of the article "Unconstrained keyword spotting using phone lattices with application to spoken document retrieval" (Foote).

In rejecting previously-presented claim 1, the Examiner stated (April 13, 2009, office action, page 6):

As per claim 1, Chou teaches:

forming a specification of a spoken event of interest to be located in unknown speech according to a plurality of sequences of subword units representing the spoken event of interest, wherein the forming includes identifying one or more instances of the spoken event of interest in a first set of audio signals and representing each identified instance of the spoken event of interest in the specification using at least one of the plurality of sequences of subword units; (Chou, column 4, lines 30-42 and column 6, lines 35-57, Fig. 2, the recognition is based on subword modeling which are compiled into networks (specification).)

The Examiner read the previously-recited "specification of a spoken event of interest" as corresponding to Chou's network of key-phrase and filler-phrase grammars. Claim 1 has been amended to make clear that a representation of a spoken event of interest is formed by "receiving an indication that a spoken event in a first set of audio signals is of interest to a user, identifying two or more instances of the spoken event of interest in the first set of audio signals, and representing each identified instance of the spoken event of interest in the representation of the spoken event of interest using at least one sequence of subword units." Chou's network of key-phrase and filler-phrase grammars is not formed as a result of such actions. Rather, Chou's network of key-phrase and filler-phrase grammars is "manually derived directly from the task specification, or, alternatively, ... generated automatically or semi-automatically (i.e., with human assistance) from a small corpus, using conventional training procedures familiar to those skilled in the art." (col. 6, lines 37-45). Even though there may be "human assistance" in the generation of Chou's network of key-phrase and filler-phrase grammars, there is no suggestion in Chou that such "human assistance" represents an "indication that a spoken event in a first set of audio signals is of interest to a user."

The Applicant agrees with the Examiner that Chou does not disclose the "accepting" and "locating" features of previously-presented claim 1. The Applicant respectfully submits that Foote does not cure these deficiencies. In rejecting

previously-presented claim 1, the Examiner stated (April 13, 2009, office action, pages 6 and 7):

Chou fails to teach, but Foote teaches:

accepting data representing the unknown speech in a second audio signal; (Foote, Page 218, ¶ 2, ...Most of the time-consuming speech recognition must be done off-line, as messages are added to the archive ... The data is input to a speech recognizer, which converts from audio to text, prior to archiving.)

locating putative instances of the spoken event of interest in the second audio signal using the specification of the spoken event of interest, wherein the locating includes identifying time locations of the second audio signal at which the spoken event of interest is likely to have occurred based on a comparison of the data representing the unknown speech with the specification of the spoken event of interest, query in the second speech data using the determined representation of the query. (Foote, Page 208, Fig. 2 and ¶ 4, ... These multiple hypotheses can be stored as a phone lattice which is a directed acyclic graph whose edges represent hypothesized phone occurrences and whose nodes represent the corresponding start and end times... Section 3.5 on Pages 214 and 215 show the keyword spotting using phone lattices.)

Foote discloses keyword spotting using a phone lattice. To do so, Foote first pre-computes a phone lattice of an unconstrained speech to be searched. This phone lattice is generated using keyword models that are derived from training data. Details about the acquisition of the training data, the content of the training data, and the derivation of the keyword models from the training data can be found on pages 210 through 216 of Foote. The phone lattice of the unconstrained speech includes multiple phone hypotheses for a set of keywords. The phone lattice is subsequently searched to find putative occurrences of a particular keyword.

On page 218, section 4.1, Foote describes both text-based and speech-based retrieval techniques:

Requests are entered as written text in natural language and common function words (such as "and", "a", "the") having little information content are removed. Once processed, a request is referred to as a search query and the words that it contains are called terms. Note that in text-based systems, the endings of query terms are usually removed using a suffix-stripping algorithm ... For speech-based retrieval, short keywords yield higher false alarm rates. Hence, suffix-stripping is less useful and it was not used here.

On page 209, section 2.3, Foote makes clear that the process of searching the phone lattice requires "a phonetic decomposition of the desired words, but these are easily found from a dictionary or by a rule-based algorithm ... "

Taken together, Foote teaches receiving a speech-based or text-based request, removing the common function words from the request, and generating a search query that is composed of a phonetic decomposition of each remaining word (i.e., each "term") of the request, where the phonetic decomposition is obtained from a dictionary or a rule-based algorithm. Accordingly, even though Foote may disclose locating putative instances of a speech-based search query in a phone lattice, Foote does not disclose or make obvious "locating ... putative instances of the spoken event of interest in the second audio signal" using a representation of the spoken event of interest that is formed by "receiving an indication that a spoken event in a first set of audio signals is of interest to a user, identifying two or more instances of the spoken event of interest in the first set of audio signals, and representing each identified instance of the spoken event of interest in the representation of the spoken event of interest using at least one sequence of subword units," as required in amended claim 1.

For at least these reasons, the Applicant respectfully submits that Chou, whether taken alone or in any proper combination with Foote, does not describe or suggest all of the features of amended claim 1.

The dependent claims 2-16 are patentable for at least similar reasons as the claims on which they depend are patentable.

The independent claims 17 and 18 are patentable for at least similar reasons given above for claim 1.

### Conclusion

It is believed that all of the pending claims have been addressed. However, the absence of a reply to a specific rejection, issue or comment does not signify agreement with or concession of that rejection, issue or comment. In addition, because the arguments made above may not be exhaustive, there may be reasons for patentability of any or all pending claims (or other claims) that have not been expressed. Finally, nothing

Applicant(s) : Robert W. Morris  
Serial No. : 10/565,570  
Filed : July 21, 2006  
Page : 11 of 11


Attorney Docket No.: 30004-004US1

in this paper should be construed as an intent to concede any issue with regard to any claim, except as specifically stated in this paper, and the amendment of any claim does not necessarily signify concession of unpatentability of the claim prior to its amendment.

No fees are believed to be due. Please apply any other charges or credits to Deposit Account No. 50-4189, referencing Attorney Docket No. 30004-004US1.

Respectfully submitted,

Date: 7/22/09

  
\_\_\_\_\_  
Mandy Jubang  
Reg. No. 45,884

Customer No. 69713  
Occhiuti Rohlicek & Tsao LLP  
10 Fawcett Street  
Cambridge, MA 02138  
Telephone: (617) 500-2530  
Facsimile: (617) 500-2499  
82930.doc